DOCUMENT RESUME

ED 174 176 HE 011 703

AUTHOR Lyons, P. R.
TITLE Program Evaluation of Graduate-Level Education. ASHE Annual Meeting 1979 Paper.
PUB DATE Apr 79
NOTE 18p.; Paper presented at the Annual Meeting of the Association for the Study of Higher Education (Washington, D.C., April 1979); Not available in paper copy due to marginal legibility of original document

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Environment; College Programs; *Educational Assessment; Educational Quality; *Evaluation Methods; Faculty Evaluation; *Graduate Study; Higher Education; *Program Evaluation; Student Evaluation
IDENTIFIERS *ASHE Annual Meeting 1979; Outcomes

ABSTRACT

The forms of evaluation of graduate-level education programs are discussed, and a process model of program evaluation which incorporates much student-generated information is presented. Graduate program reviews are generally designed either to provide information to departmental and institutional administrators regarding a program's strengths and weaknesses in a diagnostic sense or to establish the status of a particular program relative to specific performance standards or relative to performance characteristics of other similar programs. The suggested process model of graduate program evaluation relies heavily on the use of program-outcomes information. The model builds an evaluation system using quality indicators that research has indicated as having particular usefulness in the assessment of program quality. It is also based on the incorporation of evaluation methods which have been demonstrated to yield reliable and valid information for decision-making. The following aspects of the model are summarized: coordination and control; descriptive data and information needs; assessment activities for the faculty component; assessment activities for the student component; review of the program's purpose; and reporting and decision-making activities. (SC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PROGRAM EVALUATION OF GRADUATE-LEVEL EDUCATION

P. R. LYONS

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P. R. Lyons

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED174176

The purposes of this paper are: (1) to provide some general definition and scope of program evaluation of graduate-level education; (2) to report on the strengths and weaknesses of forms of evaluation of graduate programs; and (3) to present a process model of program evaluation which incorporates much student-generated types of information.

BACKGROUND

Very little research has been reported on graduate program evaluation, in general. For example, Dissertation Abstracts lists no completed dissertations on the general concept of graduate program evaluation. With the exception of some recent work completed by the Educational Testing Service there has been virtually no reported attempts of the development and implementation of general graduate program evaluation models. Much evaluation of graduate programs is conducted in-house, often with the aid of consultants. And, much evaluation activity occurs on an ad hoc basis and on a one-shot basis. A review of the available literature which is concerned with graduate program evaluation begs the conclusion that systematic, continuous graduate program evaluation is a sometimes thing in colleges and universities. Further, one may conclude, more often than not, that the need for evaluation activities are outgrowths of demands from regional accrediting agencies, professions associations, and state-level coordinating agencies and/or boards, all of which are external to the academy.

HE011 703

Over the years, there has been a growing awareness that the traditional

Association for the Study of Higher Education
The George Washington University
One Dupont Circle, Suite 630
Washington, D.C. 20036
(202) 296-2597

This paper was presented at the Annual Meeting of the Association for the Study of Higher Education held at the Washington Hilton in Washington, D.C. April 18-19, 1979. This paper was reviewed by ASHE and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC collection of ASHE conference papers.

means of evaluating graduate programs have not been without limitations. The traditional means of evaluating graduate programs have been process-oriented and component oriented relying on such measures as external consultant reviews, qualities of faculty and facilities, dollars spent, fellowships awarded, square feet assigned, research dollars attracted, number of graduates, reputational analyses, and the like. Little attention has been given to evaluating student learning, individual student change, and other outcomes aspects, in general, in colleges and universities. The National Board on Graduate Education (1 p.64) has reported that our techniques for assessing the value and quality of graduate education are in a rudimentary state. With few exceptions, it seems that graduate program evaluation has, as a practice, avoided direct confrontation with meaningful educational issues.

EVALUATION ELEMENTS

The remainder of the paper is devoted to program evaluation as it may be conducted within the college or university. There exists an uncertain consensus about the appropriate role for graduate education in our society. The higher education community may well wrestle with that issue for decades. Program administrators, however, need to make judgments about programs in the here and now hence this paper is intended to serve the practitioner's needs.

Clark (2 p.1) finds that graduate program reviews can be classified into two general categories. The first type of program review intends to provide information to departmental and institutional administrators regarding a program's strengths and weaknesses in a diagnostic sense. The purpose of such a review is to enable decision-makers to reach some consensus of the quality of the program and the needs of the program. It is anticipated that decisions for change will be made and implemented at the local level. The

second type of program review intends to establish the status of a particular program relative to specific performance standards or relative to performance characteristics of other similar programs. This latter type of review is often associated with demands for accountability and typically may be required by state-level coordinating agencies and by some accrediting groups.

The former type of review which Clark (2 p.1) has labeled "developmental" is the type which is usually of most concern and interest to the practitioner. Researchers at the Educational Testing Service (ETS) conducted a national study of assessing quality in doctoral education and their work has, among other things, identified several characteristics of programs which can be most useful in institutional research of programs. Information for the research of Clark, Hartnett, and Baird (3) was obtained from doctoral programs in chemistry, history, and psychology at 25 universities which were selected as representative of the range of all doctorate-awarding institutions in the United States. Faculty, students, recent alumni, and department chairpersons participated in the study. Three general findings of this important research have particular relevance for the purposes of this paper. First, it is possible to obtain useful and dependable information about many graduate program characteristics indicative of educational quality. Second, approximately 12-15 of these quality indicators seem to have excellent potential for use in that much information could be generated from their analysis. Finally, many of the quality indicators are applicable across disciplines which would seem to indicate that a program-wide (within an institution) evaluation system could be structured around the quality indicators.

The investigators (3 p.11.2) were able to group the indicators of program quality into six categories having to do with the quality of (1) faculty, (2) students, (3) resources, (4) environment, (5) alumni, and (6) program

academic offerings and procedures. The present investigator examined the quality indicators identified by Clark, Hartnett, and Baird (3) and applied criteria of: (1) cost of assessment, (2) applicability to all levels of graduate education, (not just doctoral-level), and (3) practicality of assessment, to the indicators. Given this analysis, several characteristics emerge as having much potential for graduate program evaluation. They are:

Faculty

1. Peer ratings of quality of faculty
   (high agreement among raters and stable over time)
2. Student rated quality of teaching
   (data easy to get)
3. Faculty - reported mean articles/book reviews in last three years
   (correlates highly with peer ratings, data easy to get)
4. Research activity
   (publications, awards, grants, etc.)

Students

5. Self-reported undergraduate grade average
   (correlates highly with faculty appraisal of scholarly ability, commitment, and motivation)
6. Student commitment, motivation
   (measured by faculty rating, correlates highly with scholarly ability and communications skills)

Resources

7. Faculty self-ratings of adequacy of facilties such as libraries, laboratories, etc.
   (data easy to collect and researchers found it to be quite reliable)

Environment

8. Student reports about various aspects of program environment, such as faculty concern for student, competitiveness among students, etc.

   (data easy to collect, is reliable, correlates with quality of teaching)

9. Faculty reports about various aspects of environment, such as program leadership, relations with other faculty in the department, etc.

   (data easy to obtain, is reliable, and may correlate highly with other quality indicators.

Program Procedures

10. Student ratings of various program elements and procedures, such as assistantship experiences, flexibility of requirements, etc.

    (data easily obtained, highly reliable)

Alumni

11. Ratings by recent alumni of various program characteristics

    (ratings are reliable, overall program quality ratings of alumni are highly correlated with self-ratings of faculty and peer ratings.

Of course, this research is based upon several very important assumptions. As Berelson (4 p. 202) pointed out nearly 20 years ago, evaluation of program quality is not a simple matter and we most often have to rely on subjective judgments regarding how good people think a particular program is. The questions persist --- what are the standards of quality and who measures them. The Task Force on Graduate Education of the Education Commission of the States (5) reports that quality considerations have been subjected to two basic

approaches in the evaluation of graduate programs. One approach calls for establishing high standards of quality and eliminating those programs which do not attain the standards. Such a process does not necessarily aid in the definition of quality. A second approach is to establish minimum standards of productivity based upon the number of degrees awarded and eliminate programs not graduating the minimum number of students. Such a process may, however, encourage poor programs to graduate lots of students.

What "Should Be" Included in the Evaluation of Graduate Programs

The Council on Postsecondary Accreditation (COPA) and the Council of Graduate Schools (COGS) issued a joint policy statement on the accreditation of graduate education (6) and in the document the organizations urge colleges and universities to be introspective and analytical in attempts at self-study and to avoid merely descriptive narratives of programs. The COPA/COGS paper goes on to say that the primary focus should be upon an assessment of outcomes, that is, "the evidence which the institution/program is producing or seeking to produce concerning the attainment of its stated objectives." (6 p. 5)

David Brenneman (7 p. 92) proposes that each graduate dean should develop a departmental information system. The system should include, as a minimum:

a. comparative costs by department;
b. Kinds/amounts of financial support available to students;
c. attrition rates and time to degree;
d. patterns of graduate placement (including where graduates are going and what they are doing);
e. procedures to monitor the quality of enrolling students.

These kinds of information are useful in the descriptive sense and can play a role in decision-making. The information however does not necessarily lend itself to assessment of outcomes.

Clark (2 p. 3) presents a checklist for graduate program review which includes the following elements:

a. a review of the program's purpose;

b. a review of the program's curricula and other educational activities (perhaps conducted by colleagues from some other campus);

c. an analysis of the usual statistical information (grades, test scores, costs, credit hours generated, etc.); and,

d. questionnaire information obtained from program graduates, faculty, and current students (primarily confined to the quality indicators represented in the research of Clark, Hartnett, and Baird (3).

In a study by Clark (8) in 1973, a group of 57 graduate school deans were utilized as an expert panel of judges concerning the aspects of graduate programs that are most clearly related to educational quality. The deans were asked to rate the importance of program characteristics from a listing of a great many program characteristics in such areas as faculty, students, resources and program operations. The deans were also asked to rate the adequacy of ways to measure the characteristics rated important. The top rated important doctoral program characteristics, which could probably be applied to any level of graduate education, were:

a. general academic ability of students;

b. achievement, knowledge, and/or skills of students at the time of degree completion;

c. financial support for program;

d. library facilities;

e. academic training of faculty;

f. purposes of the program;

g. courses and other educational offerings;

h. admissions policy; and,

i. provisions for welfare of faculty members.

With reference to an academic program as a going concern, it appears obvious that these characteristics interact continuously with one another to affect student learning and performance, and program quality. Many of the characteristics can be reliably and easily measured. Some characteristics are difficult to measure, few are fully measurable.

Two of the characteristics which seem to be central to the content of academic programs are:

(1) program purposes; and,

(2) courses and other educational offerings.

The deans reported that useful indicators for judgments about program purposes were judgments of clarity of program plans and purposes by recent graduates, faculty, or visiting experts. Clark went on to report that almost no university collected this information. The deans reported that useful indicators for judgments about courses and other educational offerings were: (1) faculty/student satisfaction with courses and other offerings; and, (2) judgments of these offerings by visiting experts. Few deans, however, reported collecting this information.

The two characteristics of graduate programs which may have the most profound impact on shaping an educational program and on the educational experiences of students are characteristics which most graduate deans identify as being extremely important --- however, they also report that information regarding these characteristics typically isn't collected. Interestingly, both characteristics rely on measurement based on asking participants or observers to report on experiences.

In a summary of the Clark et al. (3) research prepared for the Council of Graduate Schools (9 p. 18) it was recommended that a number of different assessment indicators should be incorporated for the purpose of program evaluation. Much information could be made available and if carefully arranged the information might be useful in comparisons with measures of the same characteristics of similar programs in other universities. It was also pointed out that there was no indication that program assessment respondents were yielding biased responses or otherwise systematic misrepresentation of experience. Research by Centra (10) also indicates that multiple observers, particularly students and recent graduates, yield much useful information regarding quality indicators, such as teaching performance.

A Course of Action

The research findings and the commentary of expert opinion point to a set of features which should be included in a system of program evaluation. These features are:

1. identification and articulation of program objectives;
2. use of multiple forms of assessment;
3. use of qualitative and quantitative assessment;

4. use of multiple observers of programs (students, alumni, faculty, staff, external evaluators); and
5. provision of a systematic means of information - gathering and analysis.

Research findings, expert opinion, common sense, and good administrative practice further suggest that the following features should be part of an evaluation program:

1. central coordination of evaluation activities, summaries of findings and action plans;
2. education of evaluation participants to the extent that the process is understood to be a means of improving performance; and
3. a general consensus must be achieved as to the uses to which information is to be put.

The remainder of this paper will outline the major features of the evaluation program suggested.

Coordination and Control

The evaluation program must be coordinated and controlled through the office of the dean of the graduate school (or appropriate vice-president). If there is to be a system it must be one in which decisions as to process form, exception, etc. may be centralized. This is not to say that modifications to the system cannot be accommodated, rather, the accommodation will be achieved with full knowledge of the workings of the entire evaluation program. The office of the dean should be able to manage an evaluation program with assistance from an institutional research office. The evaluation program may be conducted on a college basis with the college dean (say, in health sciences, or engineering) having been degelated certain responsibilities by the graduate dean, or the

program may be conducted with the graduate dean working directly with the department chairpersons. In any event, the graduate dean coordinates the assessment effort. The scheduling and arranging of evaluation activities are details which can be differentially treated depending on work schedules and available resources.

Descriptive Data/Information Needs

There are certain information elements which serve to describe various aspects of graduate education. These elements are easily obtained and are useful for comparative purposes. Much of the information is housed in centrally-maintained records and should be readily available. To form a body of this descriptive information, the following elements should be included:

1. data concerning the academic training of faculty;
2. comparative costs by department/academic unit;
3. statistical information by program and faculty member regarding such variables as courses taught, credit hours generated, courses cancelled, and the like;
4. kinds and amounts of financial support available to students;
5. attrition rates and time to degree; and,
6. statistical and descriptive data by program regarding such variables as undergraduate grade averages, GRE scores, and the like.

This (above) information can be accumulated in an on-going fashion independently of the information requirements for faculty and student components.

Assessment Activities for Faculty Component

The activities listed here are information-gathering ones which focus on faculty as information source. For analytical purposes, a pooling of relevant

information from several sources will be required. This includes:

1. peer ratings of quality of faculty (this can occur using internal and/or external raters, and can be conducted for all graduate faculty over a period of a few years);
2. faculty reports about various aspects of environment, such as program leadership, relations with other faculty in the department, and the like (rating scales, forced choice questions and other devices can be used to obtain this information);
3. faculty ratings of adequacy of facilities such as libraries, laboratories, and the like (information may be collected on a program basis and on an internal basis);
4. faculty self-reports of research activity (books, articles, reports published, and awards, grants, and the like); and
5. faculty reports of level of student commitment, and motivation (critical incidents and mostly subjective information will be used).

Assessment Activities for Student Component

The activities listed here are information-gathering activities which focus on students/alumni as primary information sources. The graduate dean (and other individuals) may decide on a sampling methodology to collect much of the information. Many graduate programs are so large that samples of program participants may adequately provide needed information. If personal interviews with students are part of the assessment plan, then sampling of students will be necessary. One useful way of obtaining student-generated information is to use a longitudinal approach. That is, it may be possible to identify a sample of students who begin a given program at approximately the same time and follow these students at the mid-point, and at the termination of their studies with information-gathering devices. In addition, a follow-up study of the same

students, as alumni, can be achieved a few years after the students have completed the program. The longitudinal approach has advantages of consisteny of reporting and allows for analysis of growth and development through self-reports of students. The graduate dean and other program administrators will need to carefully analyze student self-reports on the size/magnitude of reported program effects and will also need to investigate the processes that produce effects. The information provided by students will yield what has been termed information on outcomes. The information elements suggested below represent a beginning, not an ideal. The elements are:

1. student-rated quality of teaching (usually on a per-course basis);
2. reports about various aspects of program environment, such as faculty concern for student, competitiveness among students, and the like (ratings and critical incidents could be useful devices);
3. student ratings of various program elements and procedures, such as flexibility of program requirements, assistantship experiences, and the like (ratings, critical incidents, and interviews are useful information-gathering devices);
4. student self-reports of placement after program completion to include where they went and what they do; and,
5. ratings of recent alumni (2-4 years) of various program characteristics.

Program Purpose Review

Perhaps the most important element of the assessment enterprize is the review of the program's purpose. This element should receive most careful attention. Because of its importance to the evaluation effort the review of program purpose needs to be subjected to: (1) qualitative and quantitative assessments; (2) multiple observers; (3) repeated measures; and (4) repeated

consensus testing. The review may follow this example:

1. review of program's curricula and other educational activities through:
   a. assessments by outside experts;
   b. ratings, observations, etc., of faculty, students, alumni; and
   c. program profiles based upon analysis of information gathered in other measures; and,
2. the review of each program would be coordinated centrally by the dean of the graduate school with delegation of some tasks to college deans or department chairpersons, institutional research office, etc.

Reporting and Decision-Making

It is desirable for the dean of the graduate school to prepare, for internal use, a report on the program's total review every 2 or 3 years. In some cases it will be more practical to produce such a report every five years. The report will contain a program analysis based upon the information elements described previously. The report should be prepared in consultation with program directors and should be aimed at increasing the quality of the program. Specific recommendations for change should be outlined and the dean of the graduate school in conjunction with other administrators and program directors will need to reach consensus on which recommendations qualify as high-priority objectives, and how these objectives might be achieved. This specification will need to include what is to be achieved and over what period of time, who is responsible for achievement, what resources will be required for achievement, and how achievement will be evaluated. Again, this process is directed and controlled by the graduate dean.

SUMMARY

This paper presents a process-model of graduate program evaluation which relies heavily on the use of program-outcomes information. The model suggested

builds an evaluation system using thosequality indicators which research has indicated as having particular usefulness in the assessment of program quality. The model is also based upon the incorporation of evaluation methods which have been demonstrated to yield reliable and valid information for decision-making. The model provides a guide for tailoring institutional evaluation systems for program review at the graduate level.

REFERENCES

1. Outlook and Opportunities for Graduate Education. National Board on Graduate Education: Washington, D. C., Number 6, December, 1975.

2. Clark, M. J. A practical guide to graduate program review. Findings, 1979, 5 (1), 1-4.

3. Clark, M. J., Hartnett, R. T., and Baird, L. L. Assessing Dimensions of Quality in Doctoral Education: A Technical Report of a National Study in Three Fields. Princeton, N. J.: Educational Testing Service, 1976.

4. Berelson, B. Graduate education in the United States. New York: McGraw-Hill, 1960.

5. The States and Graduate Education. Report of the Task Force on Graduate Education - Education Commission of the States: Denver, Colorado, February, 1975.

6. Accreditation of Graduate Education: A Joint Policy Statement. The Council on Postsecondary Accreditation, and the Council of Graduate Schools in the United States: Washington, D. C., May, 1978.

7. Brenneman, D. W. "Decisions of the Decade." Educational Record (Summer): 220-225, 1973.

8. Clark, M. J. "The Assessment of Quality in Ph.D. Programs: A Preliminary Report on Judgments by Graduate Deans." GRE Board Research Report No. 72-7aR, Educational Testing Service, Princeton, N. J., 1973.

9. The Assessment of Quality in Graduate Education: Summary of a Multidimensional Approach. Council of Graduate Schools in the United States: Washington, D. C., 1977.

10. Centra, John A. "How Universities Evaluate Faculty Performance." GRE Board Research Report No. 75-56R, Educational Testing Service, Princeton, N.J., 1977.